EXHIBIT B

From U.S. 09/513,997
Pages 7-11, 32, 35, 36 and 37

start codon, a signal secretion sequence, an epitope tag, and a sequence-specific protease site, and an unpaired splice donor site.

The vector construct can contain one or more selectable markers for recombinant host cell selection. Alternatively, selection can be effected by phenotypic selection for a trait provided by the activated endogenous gene product.

These vectors, and indeed any of the vectors disclosed herein, and variants of the vectors that will be readily recognized by one of ordinary skill in the art, can be used in any of the methods described herein to form any of the compositions producible by these methods.

The transcriptional regulatory sequence used in the vector constructs of the invention includes, but is not limited to, a promoter. In preferred embodiments, the promoter is a viral promoter. In highly preferred embodiments, the viral promoter is the cytomegalovirus immediate early promoter. In alternative embodiments, the promoter is a cellular, non-viral promoter or inducible promoter.

The transcriptional regulatory sequence used in the vector construct of the invention may also include, but is not limited to, an enhancer. In preferred embodiments, the enhancer is a viral enhancer. In highly preferred embodiments, the viral enhancer is the cytomegalovirus immediate early enhancer. In alternative embodiments, the enhancer is a cellular non-viral enhancer.

In preferred embodiments of the methods described herein, the vector construct be, or may contain, linear RNA or DNA.

The cell containing the vector may be screened for expression of the gene. The cell over-expressing the gene can be cultured *in vitro* under conditions favoring the production, by the cell, of desired amounts of the gene product (also referred to interchangeably herein as the "expression product") of the endogenous gene that has been activated or whose expression has been increased. The expression product can then be isolated and purified to use, for example, in protein therapy or drug discovery.

Alternatively, the cell expressing the desired gene product can be allowed to express the gene product *in vivo*. In certain such aspects of the invention, the cell containing a vector construct of the invention integrated into its genome may be introduced into a eukaryote (such as a vertebrate, particularly a mammal, more particularly a human) under conditions favoring the overexpression or activation of the gene by the cell *in vivo* in the eukaryote. In related such aspects of the invention, the cell may be isolated and cloned prior to being introduced into the eukaryote.

The invention is also directed to methods for over-expressing an endogenous gene in a cell, comprising introducing a vector containing a transcriptional regulatory sequence and one or more amplifiable markers into the cell, allowing the vector to integrate into the genome of the cell by non-homologous recombination, and allowing over-expression of the endogenous gene in the cell.

The cell containing the vector may be screened for over-expression of the gene.

The cell over-expressing the gene is cultured such that amplification of the endogenous gene is obtained. The cell can then be cultured *in vitro* so as to produce desired amounts of the gene product of the amplified endogenous gene that has been activated or whose expression has been increased. The gene product can then be isolated and purified.

Alternatively, following amplification, the cell can be allowed to express the endogenous gene and produce desired amounts of the gene product *in vivo*.

It is to be understood, however, that any vector used in the methods described herein can include one or more amplifiable markers. Thereby, amplification of both the vector and the DNA of interest (i.e., containing the over-expressed gene) occurs in the cell, and further enhanced expression of the endogenous gene is obtained. Accordingly, methods can include a step in which the endogenous gene is amplified.

The invention is also directed to methods for over-expressing an endogenous gene in a cell comprising introducing a vector containing a transcriptional regulatory sequence and an unpaired splice donor sequence into the cell, allowing the vector to integrate into the genome of the cell by non-homologous recombination, and allowing over-expression of the endogenous gene in the cell.

The cell containing the vector may be screened for expression of the gene.

The cell over-expressing the gene can be cultured *in vitro* so as to produce desirable amounts of the gene product of the endogenous gene whose expression has been activated or increased. The gene product can then be isolated and purified.

Alternatively, the cell can be allowed to express the desired gene product *in vivo*.

The vector construct can consist essentially of the transcriptional regulatory sequence.

The vector construct can consist essentially of the transcriptional regulatory sequence and one or more amplifiable markers.

The vector construct can consist essentially of the transcriptional regulatory sequence and the splice donor sequence.

Any of the vector constructs of the invention can also include a secretion signal sequence. The secretion signal sequence is arranged in the construct so that it will be operably linked to the activated endogenous protein. Thereby, secretion of the protein of interest occurs in the cell, and purification of that protein is facilitated. Accordingly, methods can include a step in which the protein expression product is secreted from the cell.

The invention also encompasses cells made by any of the above methods. The invention encompasses cells containing the vector constructs, cells in which the vector constructs have integrated into the cellular genome, and cells which are over-expressing desired gene products from an endogenous gene, over-expression being driven by the introduced transcriptional regulatory sequence.

The cells can be isolated and cloned.

The methods can be carried out in any cell of eukaryotic origin, such as fungal, plant or animal. In preferred embodiments, the methods of the invention may be carried out in vertebrate cells, and particularly mammalian cells including but not limited to rat, mouse, bovine, porcine, sheep, goat and human cells, and more particularly in human cells.

A single cell made by the methods described above can over-express a single gene or more than one gene. More than one gene in a cell can be activated by the integration of a single type of construct into multiple locations in the genome. Similarly, more than one gene in a cell can be activated by the integration of multiple constructs (i.e., more than one type of construct) into multiple locations in the genome. Therefore, a cell can contain only one type of vector construct or different types of constructs, each capable of activating an endogenous gene.

The invention is also directed to methods for making the cells described above by one or more of the following: introducing one or more of the vector constructs of the invention into a cell; allowing the introduced construct(s) to integrate into the genome of the cell by non-homologous recombination; allowing over-expression of one or more endogenous genes in the cell; and isolating and cloning the cell. The invention is also directed to cells produced by such methods, which may be isolated cells.

The invention also encompasses methods for using the cells described above to over-express a gene, such as an endogenous cellular gene, that has been characterized (for example, sequenced), uncharacterized (for example, a gene whose function is known but which has not been cloned or sequenced), or a gene whose existence was, prior to over-expression, unknown. The cells can be used to produce desired amounts of an expression product *in vitro* or *in vivo*. If desired, this expression product can then be isolated and purified, for example by cell lysis or by isolation from the growth medium (as when the vector contains a secretion signal sequence).

The invention also encompasses libraries of cells made by the above described methods. A library can encompass all of the clones from a single transfection experiment or a subset of clones from a single transfection experiment. The subset can over-express the same gene or more than one gene, for example, a class of genes. The transfection can have been done with a single construct or with more than one construct.

A library can also be formed by combining all of the recombinant cells from two or more transfection experiments, by combining one or more subsets of cells from a single transfection experiment or by combining subsets of cells from separate transfection experiments. The resulting library can express the same gene, or more than one gene, for example, a class of genes. Again, in each of these individual transfections, a unique construct or more than one construct can be used.

Libraries can be formed from the same cell type or different cell types.

The invention is also directed to methods for making libraries by selecting various subsets of cells from the same or different transfection experiments.

The invention is also directed to methods of using the above-described cells or libraries of cells to over-express or activate endogenous genes, or to obtain the gene expression products of such over-expressed or activated genes. According to this aspect of the invention, the cell or library may be screened for the expression of the gene and cells that express the desired gene product may be selected. The cell can then be used to isolate or purify the gene product for subsequent use. Expression in the cell can occur by culturing the cell *in vitro*, under conditions favoring the production of the expression product of the endogenous gene by the cell, or by allowing the cell to express the gene *in vivo*.

In preferred embodiments of the invention, the methods include a process wherein the expression product is isolated or purified. In highly preferred embodiments, the cells expressing the endogenous gene product are cultured under conditions favoring production of sufficient amounts of gene product for

The present methods, therefore, are capable of identifying new genes that have been or can be missed using conventional and currently available cloning techniques. By using the constructs and methodology described herein, unknown and/or uncharacterized genes can be rapidly identified and over-expressed to produce proteins. The proteins have use as, among other things, human therapeutics and diagnostics and as targets for drug discovery.

The methods are also capable of producing over-expression of known and/or characterized genes for *in vitro* or *in vivo* protein production.

A "known" gene is directed to the level of characterization of a gene. The invention allows expression of genes that have been characterized, as well as expression of genes that have not been characterized. Different levels of characterization are possible. These include detailed characterization, such as cloning, DNA, RNA, and/or protein sequencing, and relating the regulation and function of the gene to the cloned sequence (e.g., recognition of promoter and enhancer sequences, functions of the open reading frames, introns, and the like). Characterization can be less detailed, such as having mapped a gene and related function, or having a partial amino acid or nucleotide sequence, or having purified a protein and ascertained a function. Characterization may be minimal, as when a nucleotide or amino acid sequence is known or a protein has been isolated but the function is unknown. Alternatively, a function may be known but the associated protein or nucleotide sequence is not known or is known but has not been correlated to the function. Finally, there may be no characterization in that both the existence of the gene and its function are not known. The invention allows expression of any gene at any of these or other specific degrees of characterization.

Many different proteins (also referred to herein interchangeably as "gene products" or "expression products") can be activated or over-expressed by a single activation construct and in a single set of transfections. Thus, a single cell or different cells in a set of transfectants (library) can over-express more than one protein following transfection with the same or different constructs. Previous

cellular genome in a targeted fashion, and therefore they can be used to deliver the activation construct according to the present invention.

Vector constructs useful for the methods described herein ideally may contain a transcriptional regulatory sequence that undergoes non-homologous recombination with genomic sequences in a cell to over-express an endogenous gene in that cell. The vector constructs of the invention also lack homologous targeting sequences. That is, they do not contain DNA sequences that target host cell DNA and promote homologous recombination at the target site. Thus, integration of the vector constructs of the present invention into the cellular genome occurs by non-homologous recombination, and can lead to over-expression of a cellular gene via the introduced transcriptional regulatory sequence contained on the integrated vector construct.

The invention is generally directed to methods for over-expressing an endogenous gene in a cell, comprising introducing a vector containing a transcriptional regulatory sequence into the cell, allowing the vector to integrate into the genome of the cell by non-homologous recombination, and allowing over-expression of the endogenous gene in the cell. The method does not require previous knowledge of the sequence of the endogenous gene or even of the existence of the gene. Where the sequence of the gene to be activated is known, however, the constructs can be engineered to contain the proper configuration of vector elements (e.g., location of the start codon, addition of codons present in the first exon of the endogenous gene, and the proper reading frame) to achieve maximal overexpression and/or the appropriate protein sequence.

In certain embodiments of the invention, the cell containing the vector may be screened for expression of the gene.

The cell over-expressing the gene can be cultured *in vitro* under conditions favoring the production, by the cell, of desired amounts of the gene product of the endogenous gene that has been activated or whose expression has been increased. If desired, the gene product can then be isolated or purified to use, for example, in protein therapy or drug discovery.

Alternatively, the cell expressing the desired gene product can be allowed to express the gene product *in vivo*.

The vector construct can consist essentially of the transcriptional regulatory sequence.

Alternatively, the vector construct can consist essentially of the transcriptional regulatory sequence and one or more amplifiable markers.

The invention, therefore, is also directed to methods for over-expressing an endogenous gene in a cell, comprising introducing a vector containing a transcriptional regulatory sequence and an amplifiable marker into the cell, allowing the vector to integrate into the genome of the cell by non-homologous recombination, and allowing over-expression of the endogenous gene in the cell.

The cell containing the vector is screened for over-expression of the gene.

The cell over-expressing the gene is cultured such that amplification of the endogenous gene is obtained. The cell can then be cultured *in vitro* so as to produce desired amounts of the gene product of the amplified endogenous gene that has been activated or whose expression has been increased. The gene product can then be isolated and purified.

Alternatively, following amplification, the cell can be allowed to express the endogenous gene and produce desired amounts of the gene product *in vivo*.

The vector construct can consist essentially of the transcriptional regulatory sequence and the splice donor sequence.

The invention, therefore, is also directed to methods for over-expressing an endogenous gene in a cell comprising introducing a vector containing a transcriptional regulatory sequence and an unpaired splice donor sequence into the cell, allowing the vector to integrate into the genome of the cell by non-homologous recombination, and allowing over-expression of the endogenous gene in the cell.

The cell containing the vector is screened for expression of the gene.

The cell over-expressing the gene can be cultured *in vitro* so as to produce desirable amounts of the gene product of the endogenous gene whose expression

has been activated or increased. The gene product can then be isolated and purified.

Alternatively, the cell can be allowed to express the desired gene product *in vivo*.

The vector construct can consist essentially of a transcriptional regulatory sequence operably linked to an unpaired splice donor sequence and also containing an amplifiable marker.

Other activation vectors include constructs with a transcriptional regulatory sequence and an exonic sequence containing a start codon; a transcriptional regulatory sequence and an exonic sequence containing a translational start codon and a secretion signal sequence; constructs with a transcriptional regulatory sequence and an exonic sequence containing a translation start codon, and an epitope tag; constructs containing a transcriptional regulatory sequence and an exonic sequence containing a translational start codon, a signal sequence and an epitope tag; constructs containing a transcriptional regulatory sequence and an exonic sequence with a translation start codon, a signal secretion sequence, an epitope tag, and a sequence-specific protease site. In each of the above constructs, the exon on the construct is located immediately upstream of an unpaired splice donor site.

The constructs can also contain a regulatory sequence, a selectable marker lacking a poly(A) signal, an internal ribosome entry site (ires), and an unpaired splice donor site (FIG. 4). A start codon, signal secretion sequence, epitope tag, and/or a protease cleavage site may optionally be included between the ires and the unpaired splice donor sequence. When this construct integrates upstream of a gene, the selectable marker will be efficiently expressed since a poly(A) site will be supplied by the endogenous gene. In addition the downstream gene will also be expressed since the ires will allow protein translation to initiate at the downstream open reading frame (i.e. the endogenous gene). Thus, the message produced by this activation construct will be polycistronic. The advantage of this construct is that integration events that do not occur near genes and in the proper